

091904 782  
#5

## TRANSLATION CERTIFICATION

We, **AD-EX WORLDWIDE**, an industrial translation service founded in 1957, domiciled at 525 Middlefield Road, Suite 150, Menlo Park, California 94025-3458, USA, do hereby certify that:

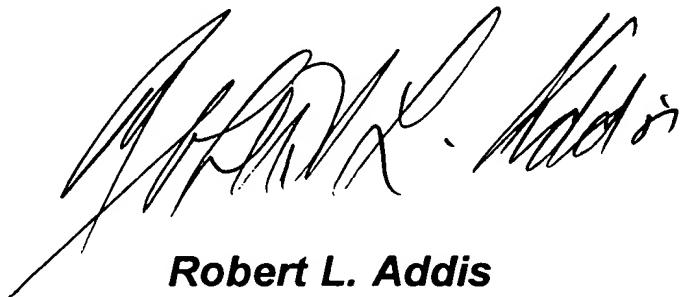
- we are internationally recognized professional translators from French into English;
- we have been serving the public in that capacity since 1957;
- the attached 19-page English text, identified on each page as AD-EX Job No. E2928A, is a translation from French to English entirely performed by us;
- said English translation from French constitutes a complete and accurate description of the entire meaning of the original French-language text identified as

***French Patent Application No. 2642767***

of which a copy is appended to this English translation.

So declared with such exceptions as may be indicated in the translation.

Thus certified on this 2nd day of May of 2000 for and on behalf of  
**AD-EX Worldwide** by its Certifying Officer:



*Robert L. Addis*

The present invention relates to new expression vectors of heterologous proteins in eukaryotic cells, in particular mammalian cells.

A heterologous protein within the meaning of the invention is understood to be a protein that is not expressed naturally by the cells in question, or else a protein that is not necessary to the survival of those cells.

Numerous vectors causing expression of genes encoding for such heterologous proteins in mammalian cells have already been described. The need to have more and more efficient vectors is constant, notably in order to appreciably increase the quantity of product synthesized by the host system and to permit the selection of productive cells and the preparation of stable cell lines producing said protein in quantities compatible with an industrial yield. Various approaches have already been suggested to improve the efficiency of this type of vector in both transcription and translation.

Thus, for example, numerous vectors use as promoter of the transcription of the DNA sequence encoding for the heterologous protein, hybrid sequences comprising the activator sequence of the SV40 virus (repeats of 72 base pairs) associated with the major late promoter of adenovirus 2. An improvement of the efficiency of this type of vector has already been obtained by various means by inserting supplemental DNA sequences originating from SV40 (repeats of 21 base pairs) and capable of binding transcription factors or tripartite leader sequences of adenovirus 2 favoring the translation of the messenger RNA.

The present invention aims to provide new vectors, useful notably for the expression of complex proteins such as the coagulation factors, in mammalian cells.

The vectors according to the invention are characterized in that the elements ensuring the expression of the sequence encoding for the heterologous protein comprise the essential elements of the leader sequence and/or of the promoter sequence of the EIII gene of adenovirus 2. An essential element is understood to be the sequence ensuring increased protein production and/or control of the expression of the sequence for the promoter.

The adenoviruses have a linear double-stranded DNA which has 35 kilobases<sup>1</sup> in the case of the serotype 2 (Ad2). The adenoviruses have five or more early transcription units, one of which is the EIII gene, and several late transcription units (1, 2). The restriction map of the upstream elements of the EIII gene is shown in Figure 1. The coordinates indicated for the nucleotides correspond to those obtained from the Ad2 sequence of the EMBL data base.

Among the important elements of the EIII gene, one distinguishes the promoter and the leader sequence, the promoter going up to the Cap site, and the leader sequence going from the Cap site to the translation initiator ATG.

Several variants of the invention can be implemented.

In a first variant, only all or part of the leader sequence of the EIII is present on the vector. Preferably, it involves a portion of the sequence which comprises an intron, so that the expression vector carries an intron in the nontranslated 5' part of the messenger RNA (see Figure 1). In that case, the promoter used can be any of the functional promoters for expressing heterologous genes in mammalian cells. For example, one can cite that of the cytomegalovirus or that of the metallothionein gene.

In a second variant, the vector according to the invention comprises all or part of the EIII gene promoter, i.e., the 5' part of the transcription elements going up to the Cap site.

A preferred variant is that in which the vector comprises simultaneously the part of the leader sequence that carries an intron and all or part of the EIII gene promoter of adenovirus 2.

It is surely preferable to have simultaneously the leader sequence and the promoter corresponding to the same gene of adenovirus 2. According to an additional characteristic of the invention, it is advantageous to add a transcription activator, in particular the activator sequences of SV40 (for example, the repeats of 72 base pairs, or else that of 21 base pairs, or an association of both).

---

<sup>1</sup> Tr. note: According to our references, Ad2 has 35,937 base pairs, i.e., closer to 36 kilobases.

In general, the vectors according to the invention are plasmids, but one can also consider viral vectors insofar as they involve viruses whose cycle comprises a step in which the genome is in the form of double-stranded DNA, e.g., the papilloma viruses. They are preferably constructed to ensure integration of the sequence encoding for the heterologous protein in the genome of the host strain, so as to obtain stable cell lines producing said protein. In order to obtain cells comprising a large number of copies of the gene encoding for the heterologous protein, the integration vectors can comprise a sequence ensuring integration of the vector in several hundred copies, such as for example the murine mitochondrial DNA fragment of the p delta plasmid.

To increase the production of the heterologous protein, it may be beneficial in some cases that the vector comprise an additional DNA sequence downstream from the DNA sequence encoding for said heterologous protein.

To select the transformed cells, the vector preferably comprises a selector gene, e.g., the gene encoding for xanthine-guanine phosphoribosyl transferase (resistance to mycophenolic acid, XGPRT), the gene for resistance to G418 (Neo), or the dihydrofolate reductase (DHFR) gene which confers resistance to methotrexate. Preferably, the Neo gene will be used.

The selector gene can be carried by an independent transcription unit, or else it can be placed directly downstream from the DNA sequence encoding for the heterologous protein. This second case is particularly advantageous because it makes it possible to achieve a double effect: increased production and improved selection efficiency.

It may also be advantageous to place the selector gene in a polycistronic expression block containing a promoter, a DNA sequence encoding for the heterologous protein, a selector gene, an intron, and a polyadenylation site. In fact, it has been demonstrated many times that ribosomes can reinitiate the downstream translation of the stop codon, and this property is used to construct plasmids containing a polycistronic selection block. As selector gene one can use those that have been cited above; the mouse gene encoding for dihydrofolate [sic] reductase (DHFR) and the bacterial gene of XGPRT.

In order to be able to more easily select, after transformation of mammalian cells, those which express the heterologous protein in quantity, it is particularly advantageous to add to a polycistronic unit a transcription unit of a selector gene. Preferably, an expression block of the DHFR gene will be used, e.g., under the control of the early promoter of SV40 deprived of its activator sequence.

The invention also has as its object a process for preparation of eukaryotic cells in order to obtain cell lines producing a heterologous protein, according to which the transformation of these cells is brought about by a vector according to the invention, and the cells expressing said proteins<sup>2</sup> are selected.

Among the transformation methods used, one may cite in particular the technique of calcium phosphate coprecipitation which will be implemented in the examples hereinbelow.

Finally, the invention has as its object the cell lines obtained by implementation of the process according to the invention, in particular the CHO [*Chinese hamster ovary*] lines producing Factor VIII or Factor IX. In fact, the invention is applicable to the preparation of blood coagulation factors and hence also concerns a process for their preparation, according to which the cells are cultivated on an appropriate culture medium and the protein obtained is recovered from the culture.

Other characteristics and advantages of the invention will appear in the course of the detailed description to follow, which will indicate some examples of constructions according to the invention and their application to the expression of complex proteins in mammalian cells.

The following Figures illustrate the invention:

- Figure 1 schematizes the restriction map of the upstream transcription elements of the EIII gene.

---

<sup>2</sup> *Tr. note: The identically worded claim 16 has "said heterologous protein" [singular], which makes more sense.*

- Figure 2 shows schematically the structure of the pSB394 plasmid.
- Figure 3 shows schematically the structure of the pSBDH394 plasmid.
- Figure 4 shows schematically the structure of the pTGEIII385, pTGEIIDH385 and pTGEIIINeo385 plasmids.
- Figure 5 shows schematically the structure of the pTG2307 plasmid.
- Figure 6 shows schematically the structure of the pTG2330 plasmid.

Example 1                    Transient expression of Factor IX in CHO cells

a)        Construction of pSB394 and pSBDH394

The pSB394 plasmid (Fig. 2) is a vector derived from pTG157 described in the European patent publication EP.A.0140762 where the EcoRI - BglII sequences encoding for the rabic glycoprotein have been eliminated and replaced by the following fragments:

- the EcoRI - KpnI fragment of the polylinker of M13TG131 (3);
- the entire activator of the SV40 virus in the form of the KpnI - HindII fragment of pSO (4);
- a part of the late major promoter sequence of adenovirus 2, SmaI - BamHI of pM4 (5);
- the BamHI - BglII fragment of the polylinker of M13TG127, a vector comprising a polylinker with the sequence  
5' AGATCTGCAGGTCCAAGCTTGGACGGATCCCCGGGGA?TTC 3' comprising the BglII, PstI, HindIII, BamHI, SmaI and EcoRI restriction sites;

- the complementary DNA of Factor IX BamHI<sup>3</sup> of M13TG315 inserted into the BglII site of the polylinker. It is obtained by digesting the M13MP701 vector (3) by the BamHI and EcoRI enzymes, then inserting the EcoRI - PstI sequences of M13MP8 (6), the large PstI - FnuDII fragment of the complementary DNA of Factor IX (of pTG397 (7)) and an FnuDII - BamHI adaptor with the sequence

5' GATCCATGCAGCG 3'

3' GTACGTCGC 5'

restoring the complete encoding sequence of the complementary DNA.

The pSB394 and pSBDH394 plasmids differ first of all in the integration site of the complementary DNA of Factor IX. While in pSB394 this DNA is inserted into the BglII site of the polylinker, in pSBDH394 it is into the BamHI site that it is inserted.<sup>4</sup> Moreover, in a second step the complementary DNA of murine DHFR (HindIII - BglII fragment of pMTVdhfr (8)) is introduced between the HindIII and BglII sites of the polylinker (see Fig. 3).

b) Construction of pTGEIII385, pTGEIIDH385 and pTGEIIINeo385 plasmids containing a fragment of the EIII gene of adenovirus 2

To implement these constructions, one can use adenovirus 2 deposited under ATCC No. VR-846. The pTGEIII385 plasmid (Fig. 4) is obtained by digesting pSB394 by EcoRI and BamHI, then replacing the late major promoter of adenovirus 2 and the cDNA of Factor IX by the following sequences:

- the EcoRI - SmaI fragment of adenovirus 2 between the sequences 27372 and 27571 (see Fig. 1) (the coordinates correspond to those obtained from the Ad2 sequence of the EMBL data base);

---

<sup>3</sup> Tr. note: Something appears to be missing or scrambled in the preceding few words.

<sup>4</sup> Tr. note: The French preposition "dans", which appears 4 times in this sentence alone, can be translated as either "in" or "into" in English. Which is correct depends entirely on the context. Since the object of the preposition is often just a symbol composed of letters and numbers, we cannot guarantee that all our best guesses will be correct.

- the EcoRV - KpnI fragment of the polylinker of M13TG131 (3);
- the activator of the SV40 virus in the form of the KpnI - HindII fragment of pSO (4);
- the adenovirus 2 sequence from nucleotide 27572 to 28812 (from the SmaI site of the EIII promoter to the ATG; see Fig. 1);
- the complementary DNA of Factor IX fused to the initiator ATG of EIII.

To implement this fusion of the complementary DNA of Factor IX and the EIII gene at the initiator ATGs, the complementary DNA of Factor IX is excised in the form of the BamHI-XhoII fragment by partial digestion of M13TG390 (described below) and cloned in the BamHI site of M13TG131 giving the M13TG373 vector.

M13TG390 derives from M13TG120 (3) and contains the following sequences between the SalI and PstI sites:

- an SalI - FnuDII adaptor having the sequence



- The FnuDII - PstI fragment of the cDNA of factor IX coming from pTG397 described in the European patent publication EP-167420.

The HindIII fragment of the leader of the EIII gene (nucleotides 28653 to 28962, see Fig. 1) is inserted into the HindIII site of M13TG373 and a loop mutagenesis<sup>5</sup> using an oligonucleotide with the sequence

5' GGGCAACATCCAAGATGCAGCGCGTGAACATGATC 3' makes it possible to obtain fusion of the complementary DNAs of Factor IX and the EIII gene at the initiator ATGs.

---

<sup>5</sup> Tr. note: Presumably meant to be the same as the "loop-deletion mutagenesis" appearing several pages further on, i.e., production of a mutation comprising a deletion loop in a DNA strand.

The pTGEIIINeo385 plasmid (Fig. 4) results from insertion of the BamHI fragment of M13TG1823 into the BamHI site of pTGEIII385. This vector derives from M13MP9 (6); in its BamHI site it contains a fragment of the neo gene obtained by BamHI-BglII digestion of pSv2-neo (9). Once introduced, this fragment is mutated locally with an oligonucleotide having the sequence 5' CATGCGAAAGGATCCTCATCC 3', making it possible to excise the neo gene in the form of a BamHI fragment of M13TG1823.

The pTGEIIIDH385 plasmid (Fig. 4) is obtained by inserting the DHFR gene fragment of pMTVdhfr (8) into the HindIII and BglII sites of pTGEIII385.

c) Transient expression in CHO cells

Culture dishes (5 cm diameter) are seeded with 300,000 cells. At the end of 24 hours, the DNA of the expression vectors of Factor IX (2 µg) and the DNA of an expression vector of beta galactosidase, pCH110 (10) (2 µg), are cotransferred by means of the standard technique of calcium phosphate coprecipitation (11). After 24 hours, the cells are washed with PBS (137 mM NaCl; 2.7 mM KCl; 8 mM Na<sub>2</sub>HPO<sub>4</sub>; 1.5 mM KH<sub>2</sub>PO<sub>4</sub>; pH = 7.4) and new medium is added. After a period of 48 hours, the culture media are prewashed and the quantity of Factor IX produced, F, is assayed by means of an ELISA assay marketed by Diagnostica-Stago. The cells are washed 3 times with PBS, then centrifuged. They are recovered in 0.2 to 1 ml of buffer Z (10) and ruptured by means of brief ultrasonication. The cell debris is eliminated by means of centrifugation; the beta galactosidase activity B is assayed in the supernatant by a standard method (12) and the total quantity of proteins P is determined by a Biorad assay. The efficiency of the constructions is evaluated via the ratio F·P/B.

The plasmids of the pSB series (pSB394 and pSBDH394) use promoters containing sequences of the late major promoter of adenovirus 2 joined to sequences of the SV40 activator, while the plasmids of the pTGEIII series (pTGEIIIDH385, pTGEIIINeo385 and pTGEIII385) according to one aspect of the present invention comprise the leader sequence and a part of the promoter of the EIII gene of adenovirus 2 as well as the SV40 activator sequences upstream of the promoter.

pSB394 and pTGEIII385 are two monocistronic vectors, while pSBDH394, pTGEIIDH385 and pTGEIIINeo385 are polycistronic, using as second gene the cDNA of the DHFR gene for the first two and the neo gene for the last.

The results of transient expression of Factor IX in CHO cells are shown in Table I below.

**Table I:** Measurement of the transient expression of Factor IX in CHO cells transfected by different plasmids

<u>Plasmid</u>	<u>F x p/B *</u>
pSB394	2.8
pSBDH394	7.8
pTGEIII385	17.0
pTGEIIDH385	105.0
pTGEIIINeo385	150.0

\* B:  $\beta$ -galactosidase activity

P: total quantity of proteins

F: Factor IX activity measured by ELISA

Table I demonstrates the following results:

- 1) Replacement of the late major promoter of adenovirus 2 by the elements of the EIII gene of adenovirus 2 in the starting plasmid makes it possible to improve the quantity of Factor IX produced by a factor of 5 to 12.
- 2) Addition of a sequence downstream of the cDNA of factor IX improves the efficiency of the constructions (pTGEIII385/pTGEIIDH385).
- 3) Addition of the sequence of the neo gene downstream of the cDNA of Factor IX gives a slightly better result than addition of the sequence of the cDNA of the DHFR gene (pTGEIIDH385/pTGEIIINeo385).

Example 2: Expression of Factor VIIIΔII

The analog of Factor VIII, designated by FVIIIΔII, has been described in the French patent application No. 8711415 filed in the applicant's name. It is characterized in that it involves the compound in which amino acids 771 to 1662 have been deleted.

a) Construction of pTG2307

The pTG2307 plasmid (Fig. 5) contains a part of the p delta plasmid favoring integration of the vector in a large number of copies in the genome and a polycistronic transcription unit (promoter + leader) EIII - Factor VIIIΔII/XGPRT.

It is constructed as follows:

The pTG2307 plasmid derives from pTG1509 (deposited in the CNCM under No. I-681 on 24 July 1987) described in the previously cited French patent application No. 8711415 digested by the NcoI and Aval restriction enzymes. The short fragment is eliminated and replaced by the PvuI - Aval fragment of M13TG347 (described below) with the aid of an NcoI - PvuI adaptor having the sequence:

5' CCGGCCTAGGCCGGCTGCAGAT 3'  
3' GGATCCGGGCCCCGACGTC 5'

M13TG347 derives from M13TGE131 (3) by insertion of the following fragments into the EcoRI and KpnI sites:

- EcoRI - EcoRV of pTGEIII385 carrying the EIII promoter followed by its leader and a part of the cDNA of Factor IX
- SmaI - KpnI of pTG1080 (13) carrying the 5' end of the cDNA of Factor VIII.

Fusion of the EIII proteins and Factor VIII at the ATGs is implemented by means of a deletion-loop mutagenesis using an oligonucleotide having the sequence

5' CTGGGGCAACATCCAAGAATGCAAATAGAACTCTCC 3'.

b) Production of Factor VIIIΔII in CHO cells

The pTG2307 vector thus obtained is transferred into the CHODXB11 cells (Dr. L. Chasin, Columbia University, New York, USA) by the calcium phosphate coprecipitation technique and the transformed cells are selected by the process described in Reference (14), but omitting the aminopterin. The clones are isolated after 1.5 to 2 months of selection, then tested for their production of Factor VIIIΔII. Among 14 clones tested, 7 produce between 1 and 6 U of Factor VIII (coagulating activity) per  $10^6$  cells per 24 hours. These production levels are 12 times greater than those obtained using the SV40 promoter of adenovirus 2 of pTG384 (construction pTG1509, described in the previously cited French patent application No. 8711415) (see Table II).

Table II: Production of Factor VIIIΔII in CHO cells as a function of the plasmid used

Plasmid	Coagulating activity U / $10^6$ cells x 24 h
pTG1509	0.5
pTG2307	6.0

Example 3: Production of von Willebrand factor with the aid of the pTG2330 plasmid (Fig. 6)

The starting plasmid is the p delta plasmid, more precisely the large fragment of p delta between the EcoRI and SalI sites. The promoter and the leader of the EIII gene are introduced as EcoRI - HindIII fragment of pTGEIII385, then a HindIII - BamHI adaptor coming from M13TG131 (3) and the block of the Neo gene [*in the form of a*] BamHI fragment of M13TG1823 (described previously) bound to the BglII - SalI fragment of pTGEIII385. The plasmid thus formed is designated by pTG2323.

To obtain the DHFR block, the pSBDH394 plasmid is digested by HindIII and PvuII, then the Factor IX expression block is replaced by the SV40 early promoter, HindII - HindIII fragment of p delta. The HindIII site is then eliminated: the plasmid is digested by HindIII, the site is filled by DNA polymerase (Klenow fragment) and the plasmid is religated.

The DHFR expression block is excised by SalI - SphI digestion and cloned in M13TG131 (3) cut by those same enzymes, yielding M13TG346. Thus the DHFR block is understood to be a transcription unit formed of an expression block of the DHFR gene under the control of the SV40 early promoter deprived of its activator sequence. The pTG2323 plasmid is then digested by EcoRI, then treated with S1 nuclease. The ends are replaced with DNA polymerase (Klenow fragment), then ligated with the DHFR block, SmaI - SalI fragment of M13TG346 (SalI end filled with Klenow). Fusion of the SalI site filled with Klenow and the EcoRI site treated with S1 nuclease regenerates an SalI site.

The DHFR block is integrated in such a manner that the EIII and DHFR genes are transcribed in opposite directions (pTG2325). A polylinker that has the sequence



and that has the HindIII, NruI, EcoRI, XhoI and BamHI restriction sites is introduced between the HindIII and BamHI sites of pTG2325, yielding pTG2328. The polylinker makes it possible to insert the cDNAs that are to be expressed. The expression vector of the von Willebrand (vWF) factor is obtained by inserting the cDNA in the form of an EcoRI fragment into the EcoRI site of pTG2328, yielding pTG2330.

The cDNA encoding for complete vWF is prepared from mRNA extracted from human lung, an organ rich in vascular tissue and thus in endothelial cells which are recognized as a vWF synthesis site. The cDNA is cloned in a *gt10* bacteriophage. The clones obtained are screened using nucleotide strains (13).

Several clones were obtained and then assembled so as to obtain a cDNA clone encoding for the entire protein. The plasmid containing the assembled sequences is composed of the pTZ18R vector (Pharmacia) containing in its EcoRI site the cDNA of the vWF from nucleotide -131 (assuming that the initiator ATG is at +1) to nucleotide 8557 (SstI site). By construction, the cDNA comprises a polyG end at its 5' end and an SstI-EcoRI fragment of M13TG131 (3) at its 3' end.

The pTG2330 vector contains two transcription units: a vWF/neo polycistronic block and the DHFR block. Consequently, it is possible to select the transformed clones by using one of the neo or DHFR markers or both of them simultaneously.

Thus, the pTG2330 vector has been transfected into the CHO-DHFR<sup>-</sup> DG44 line (obtained from Dr. L. Chasin, Columbia University, New York, USA) using the calcium phosphate precipitation [*sic*] method. Three days after the transfer, the cells are trypsinated and distributed in an Alha [*sic*] 1900 medium (GIBCD) supplemented with 10% fetal calf serum (FCS) and with G418 antibiotic at a concentration of 400 µg/ml or in an Alpha 2000 medium supplemented with FCS, with (at a concentration of 10 nM or 30 nM) or without methotrexate. After two days of selection, the cells in G418 are trypsinated and distributed either in an Alpha 1900 medium supplemented with FCS and containing 600 µg/ml of G418 or in an Alpha 2000 medium (GIBCO) supplemented with dialyzed FCS and containing 200 or 400 µg/ml of G418. The cells are kept in these media for two to three weeks; then the clones that have emerged are tested for their production of vWF by means of an ELISA assay marketed by Diagnostic-Stago.

Productive clones can be obtained under all selection conditions (Table I). The most productive ones are isolated in Alpha 2000 supplemented with methotrexate.

At this stage, it may be of interest to improve the transfer conditions in the cells so as to obtain clones that are capable of resisting higher concentrations of methotrexate, or to be able to derive lines resistant to higher concentrations of methotrexate from clones that have already been isolated.

**Table III** Production of CHO-DG44 clones transformed by pTG2330 under the selection conditions

Selection medium	Number of clones obtained having a vWF production in ng/10 <sup>6</sup> cells/24 h) of:				
	0-10	10-100	100-500	500-1000	1000-3000
Alpha 2000 FCS	26	8	20	1	0
Alpha 2000 FCS Methotrexate 10 nM	5	0	3	3	4
Alpha 2000 FCS Methotrexate 10 nM	1	0	0	0	2

## REFERENCES

- (1) RIGBY, F.W.J. in Genetic Engineering Vol. 3, edited by R. WILLIAMSON. Academic Press, pp. 88-141.
- (2) ZIFF, F. (1980). Nature 287, 491-499.
- (3) KIENY, M.P., LATHE, R. and LECOCQ, J.P. (1983). Gene 26, 91-99.
- (4) TAHASASHI, K. et al. (1986). Nature 319, 121-126.
- (5) MIYAMOTO, N.G. et al. (1984). Nucl. Acids Res. 12, 8779-8799.
- (6) MESSING, J. and VIEIRA, J. (1982). Gene 19, 269-276.
- (7) JAYE, M. et al. (1983). Nucl. Acids Res. 8, 2325-2335.
- (8) LEE, F. et al. (1981). Nature 294, 228-232.
- (9) SOUTHERN, P.J. and BERG, P. (1982). J. Mol. Appl. Genet. 1, 327-341.
- (10) HALL, C.V. et al. (1983). J. Mol. Appl. Genet. 2, 101-109.
- (11) GRAHAM, F.L. and A.J. VAN DER EB (1973). Virol. 52, 456-467.
- (12) MILLER, J.H. (1972) in Experiments in molecular genetics. Cold Spring Harbor.
- (13) PAVIRANI, A. et al. (1987). Biotechnology 5, 389-392.
- (14) MULLIGAN, R.C. and BERG, P. (1981). Proc. Natl. Acad. Sci. USA 78, 2072-2076.
- (15) WU Q.Y. et al. (1988). Blood 71, 1341-1346.

## CLAIMS

1. Expression vector of a heterologous protein in the eukaryotic cells, comprising a DNA sequence encoding for the heterologous protein as well as the elements ensuring its expression, characterized in that the elements ensuring the expression of the sequence encoding for the heterologous protein comprise the essential elements of the leader sequence and/or of the promoter sequence of the EIII gene of adenovirus 2.
2. Vector according to claim 1, characterized in that it comprises all or part of the leader sequence of the EIII gene.
3. Vector according to claim 2, characterized in that the part of the leader sequence of the EIII gene is that which comprises the intron (nucleotides 27981-28360).
4. Vector according to claim 3, characterized in that the part of the leader sequence of the EIII gene is that which begins at the intron (nucleotide 27981) and extends up to the translation initiator ATG.
5. Vector according to claim 1, characterized in that it comprises all or part of the promoter of the EIII gene.
6. Vector according to any of claims 1 to 5, characterized in that it comprises all or part of the promoter and the part of the leader sequence that comprises the intron (nucleotides 27981-28360) of adenovirus 2.
7. Vector according to claim 6, characterized in that it comprises additionally a transcriptional activator sequence.
8. Vector according to any of the preceding claims, characterized in that it comprises elements ensuring its integration into the genome of the host cell, preferably the murine mitochondrial DNA sequence of the p delta plasmid.

9. Vector according to any of claims 1 to 8, characterized in that it comprises an additional DNA sequence downstream of the DNA sequence encoding for the heterologous protein.
10. Vector according to any of the preceding claims, characterized in that it comprises a selector gene.
11. Vector according to claim 10, characterized in that the selector gene is located downstream of the DNA sequence encoding for the heterologous protein.
12. Vector according to claim 11, characterized in that the selector gene is located on a polycistronic expression block downstream of the DNA sequence encoding for the heterologous protein.
13. Vector according to claim 11 or 12, characterized in that it involves the Neo gene.
14. Vector according to any of claims 10 to 13, characterized in that it comprises a transcription unit of a first selector gene and a polycistronic expression block for the heterologous protein comprising a second selector gene.
15. Vector according to claim 14, *[characterized]* in that the first selector gene is the DHFR gene.
16. Process for preparation of eukaryotic cells in order to obtain cell lines producing a heterologous protein, characterized in that the transformation of these cells is brought about by a vector according to claims 1 to 15, and the cells expressing said heterologous protein are selected.
17. Process according to claim 16, characterized in that the transfection is performed by the calcium phosphate coprecipitation technique.
18. Cell line obtained by implementation of the process according to any of claims 16 or 17.

19. Process for preparation of a heterologous protein by cell lines, characterized in that a cell line according to claim 18 is cultured on an appropriate culture medium, and the heterologous protein is recovered from the culture.
20. Process according to claim 19, characterized in that the heterologous protein is chosen from among the blood coagulation factors.

Key for French words in Figures 1 to 6 appended to the patent text

promoteur = promotor  
facteur = factor

(19) FRENCH REPUBLIC  
 NATIONAL INSTITUTE OF  
 INDUSTRIAL PROPERTY  
 PARIS

(11) Publication N.: **2 642 767**  
 (To be used only for  
 reproduction orders)

(21) National Registration N.: **89 00599**

(51) Int. Cl.<sup>5</sup>: C 12 N 15/67, 15/86.

(12)

## PATENT APPLICATION

**A1**

(22) Filing date:	19 January 1989	(71) Applicant(s): <i>TRANSGENE S.A., Limited liability company - FR.</i>
(30) Priority:		(72) Inventor(s): <i>Henri de La Salle; Thérèse Faure.</i>
(43) Date on which application was made available to the public:	<i>BOP - "Patents", No. 32 of 10 August 1990.</i>	(73) Patentee(s):
(60) References to other related national documents:		(74) Agent(s): <i>Cabinet Regimbeau, Martin, Schrimpf, Warcoin et Ahner</i>

(54) Expression vectors of heterologous proteins in eukaryotic cells, cell lines obtained and process for their preparation.

(57) The present invention relates to an expression vector of a heterologous protein in the eukaryotic cells, comprising a DNA sequence encoding for the heterologous protein as well as the elements ensuring its expression, characterized in that the elements ensuring the expression of the sequence encoding for the heterologous protein comprise the essential elements of the leader sequence and/or of the promoter sequence of the E111 gene of adenovirus 2.

---

**FR 2 642 767 - A1**

- La figure 5 représente schématiquement la structure du plasmide pTG2307.

5 - La figure 6 représente schématiquement la structure du plasmide pTG2330.

Exemple 1 Expression transitoire de facteur IX dans des cellules CHO

10

a) Construction de pSB394 et pSBDH394

Le plasmide pSB394 (fig. 2) est un vecteur dérivé de pTG157 décrit dans la publication européenne de brevet EP.A.0140762 où les séquences EcoRI - BglII codant pour la glycoprotéine rabique ont été éliminées et remplacées par les fragments suivants :

20 - le fragment EcoRI - KpnI du polylinker de M13TG131 (3)

25 - l'activateur entier du virus SV40 sous forme du fragment KpnI - HindII de pS0 (4) ;

30 - une partie de la séquence du promoteur majeur tardif de l'Adénovirus 2, SmaI - BamHI de pM4 (5) ;

35 - le fragment BamHI - BglII du polylinker de M13TG127, vecteur comprenant un polylinker de séquence 5' AGATCTGCAGGTCCAAGCTTGGACGGATCCCCGGGG 3' comportant des sites de restriction BglII, PstI, HindIII, BamHI, SmaI et EcoRI.

- l'ADN complémentaire du facteur IX BamHI de M13TG315 inséré dans le site BglII du polylinker. Il est obtenu en digérant le vecteur M13MP701 (3) par les enzymes

BamHI et EcoRI puis en insérant les séquences EcoRI - PstI de M13MP8 (6), le grand fragment PstI - FnuDII de l'ADN complémentaire du facteur IX (de pTG397 (7)) et un adaptateur FnuDII - BamHI de séquence

5' GATCCATGCAGCG 3'

3' GTACGTCGC 5'

restaurant la séquence codante complète de l'ADN complémentaire.

10 Les plasmides pSB394 et pSBDH394 diffèrent en premier lieu par le site d'intégration de l'ADN complémentaire du facteur IX. Si dans pSB394 cet ADN est inséré dans le site BglII du polylinker, dans pSBDH394, c'est dans le site BamHI qu'il est inséré. De plus, en 15 une seconde étape, l'ADN complémentaire de DHFR de souris (fragment HindIII - BglII de pMTVdhfr(8)) est introduit entre les sites HindIII et BglII du polylinker (voir fig. 3).

20 b) Construction des plasmides pTGEIII385, pTGEIIIDH385 et pTGEIIINeo385 contenant un fragment du gène EIII de l'adénovirus 2

Pour réaliser ces constructions, on peut utiliser 25 l'adénovirus 2 déposé sous N° ATCC VR-846. Le plasmide pTGEIII385 (fig. 4) est obtenu par digestion de pSB394 par EcoRI et BamHI puis en substituant le promoteur majeur tardif de l'adénovirus 2 et l'ADNc du Facteur IX par les séquences suivantes :

30 - le fragment EcoRI - SmaI de l'Adénovirus 2 entre les séquences 27372 et 27571 (voir fig. 1) (les coordonnées correspondent à celles obtenues à partir de la séquence Ad2 de la banque de données de l'EMBL) ;

35 - le fragment EcoRV - KpnI de polylinker de M13TG131 (3) ;

- l'activateur du virus SV40 sous forme du fragment KpnI - HindIII de pSO (4) ;
- la séquence de l'Adénovirus 2 des nucléotides 27572 à 5 28812 (du site SmaI du promoteur de EIII à l'ATG, voir fig. 1).
- l'ADN complémentaire du facteur IX fusionné à l'ATG initiateur de EIII.

10 Pour réaliser cette fusion de l'ADN complémentaire du facteur IX et du gène EIII au niveau des ATG initiateurs, l'ADN complémentaire du facteur IX est excisé sous forme du fragment BamHI-XbaII par 15 digestion partielle de M13TG390 (décris ci-après) et cloné dans le site BamHI de M13TG131 donnant le vecteur M13TG373.

20 M13TG390 dérive de M13TG120 (3) et contient entre les sites SalI et PstI les séquences suivantes :

- un adaptateur SalI - FnuDII de séquence :

25 5' TCGACCATGCAGCG 3'  
3' GTACGTCGC 5'
- le fragment FnuDII - PstI de l'ADNc du facteur IX provenant de pTG397 décris dans la publication européenne de brevet EP-167420.

30 Le fragment HindIII du leader du gène EIII (nucléotide 28653 à 28962, voir fig. 1) est inséré dans le site HindIII de M13TG373 et une mutagénèse par boucle utilisant un oligonucléotide de séquence  
35 5' GGGGCAACATCCAAGATGCAGCGCGTGAACATGATC 3' permet

d'obtenir la fusion des ADN complémentaires du facteur IX et du gène EIII au niveau des ATG initiateurs.

Le plasmide pTGEIIINéo385 (fig. 4) résulte de  
5 l'insertion dans le site BamHI de pTGEIII385 du fragment  
BamHI de M13TG1823. Ce vecteur dérive de M13MP9 (6) ; il  
contient dans son site BamHI un fragment du gène néo,  
obtenu par digestion BamHI-BglIII de pSv2-néo(9). Une fois  
introduit, ce fragment est muté localement avec un  
10 oligonucléotide de séquence 5'CATGCGAAAGGATCCTCATCC 3'  
permettant d'exciser le gène néo sous forme d'un fragment  
BamHI de M13TG1823.

Le plasmide pTGEIIIDH385 (fig. 4) est obtenu par  
15 insertion dans les sites HindIII et BglIII de pTGEIII385  
du fragment de gène DHFR de pMTVdhfr (8).

c) Expression transitoire dans les cellules CHO

20 Des boîtes de culture (diamètre 5 cm) sont  
ensemencées avec 300 000 cellules. Au bout de 24 heures  
l'ADN des vecteurs d'expression du facteur IX (2 µg) et  
l'ADN d'un vecteur d'expression de la beta-galactosidase,  
pCH110 (10), (2 µg) sont cotransférés selon la technique  
25 classique de coprécipitation au phosphate de calcium  
(11). Après 24 heures, les cellules sont lavées avec PBS  
[137 mM NaCl ; 2,7 mM KCl ; 8 mM Na<sub>2</sub>HPO<sub>4</sub> ; 1,5 mM KH<sub>2</sub>PO<sub>4</sub> ;  
pH = 7,] et du milieu neuf est ajouté. Après un délai de  
48 heures les milieux de culture sont prélevés et la  
30 quantité de facteur IX produit, F, est dosée au moyen  
d'un test ELISA commercialisé par Diagnostica-Stago. Les  
cellules sont lavées 3 fois avec du PBS puis centrifugées.  
Elles sont reprises dans 0,2 à 1 ml de tampon Z  
(10) et cassées au moyen d'une courte sonication. Les  
35 débris cellulaires sont éliminés par centrifugation,

l'activité beta-galactosidase, B, est dosée dans le surnageant suivant une méthode classique (12) et la quantité totale de protéines, P, est déterminée par test Biorad. L'efficacité des constructions est évaluée par le rapport F.P/B.

Les plasmides de la série pSB (pSB394 et pSBDH394) utilisent des promoteurs contenant des séquences du promoteur majeur tardif de l'Adénovirus 2 jointes aux séquences de l'activateur de SV40 alors que les plasmides de la série pTGEIII (pTGEIIIDH385, pTGEIIINeo385 pTGEIII385) conformément à un aspect de la présente invention comportent la séquence leader et une partie du promoteur du gène EIII de l'adénovirus 2 ainsi que en amont du promoteur les séquences activatrices de SV40.

pSB394 et pTGEIII385 sont deux vecteurs monocistroniques alors que pSBDH394, pTGEIIIDH385 et pTGEIIINeo385 sont polycistroniques utilisant comme second gène, l'ADNc du gène DHFR pour les deux premiers et le gène neo pour le dernier.

Les résultats d'expression transitoire du Facteur IX dans les cellules CHO sont exposés dans le tableau I ci-après.

Tableau I : Mesure de l'expression transitoire du Facteur IX dans des cellules CHO transfectées par différents plasmides.

12

14

	<u>Plasmide</u>	<u>F x P*</u>
	B	
	pSB394	2.8
	psBDH394	7.8
5	pTGEIII385	17.0
	pTGEIIIDH385	105.0
	pTGEIIINeo385	150.0

\* B : activité β-galactosidase

10 P : quantité totale de protéines

F : activité du facteur IX mesurée par ELISA

Le tableau I met en évidence les résultats suivants :

15 1) Le fait de remplacer dans le plasmide de départ, le promoteur majeur tardif de l'adénovirus 2 par les éléments du gène EIII de l'adénovirus 2, permet d'améliorer d'un coefficient 5 à 12 la quantité de facteur IX produite.

20 2) L'adjonction d'une séquence en aval de l'ADNc du facteur IX améliore d'un facteur 5 l'efficacité des constructions (pTGEIII385/pTGEIIIDH385).

25 3) L'adjonction de la séquence du gène néo en aval de l'ADNc du facteur IX donne un résultat légèrement meilleur que l'adjonction de la séquence de l'ADNc du gène DHFR (pTGEIIIDH385/pTGEIIINeo385).

30 Exemple 2 : Expression du facteur VIIIΔII

L'analogique du facteur VIII désigné par FVIIIΔII a été décrit dans la demande française de brevet n° 8711415 déposée au nom de la Demanderesse. Il est caractérisé en

ce qu'il s'agit du composé déléte des acides aminés 771 à 1662.

a) Construction de pTG2307

5

Le plasmide pTG2307 (fig. 5) contient une partie du plasmide p delta favorisant l'intégration du vecteur en un grand nombre de copies dans le génome et une unité de transcription polycistronique (promoteur + leader) EIII - Facteur VIIIΔIII/XGPRT.

10

Il est construit de la façon suivante :

15

Le plasmide pTG2307 dérive de pTG1509 (déposé à la CNCM sous le numéro I-681 le 24 juillet 1987) décrit dans la demande française de brevet déjà citée n° 8711415 digéré par les enzymes de restriction NcoI et AvaI. Le fragment court est éliminé et substitué par le fragment PvuI - AvaI de M13TG347 (décrit ci-après) avec l'aide d'un adaptateur NcoI - PvuI de séquence :

20

5' CCGGCCTAGGCCGGGCTGCAGAT 3'  
3' GGATCCGGGCCCCGACGTC 5'

25

M13TG347 dérive de M13TG131 (3) par insertion dans les sites EcoRI et KpnI des fragments :

- EcoRI - EcoRV de pTGEIII385 portant le promoteur EIII suivi de son leader et une partie de l'ADNc du facteur IX.

30

- SmaI - KpnI de pTG1080 (13) portant l'extrémité 5' de l'ADNc du facteur VIII.

35

La fusion au niveau des ATG des protéines EIII et du facteur VIII est réalisée par une mutagénèse par délétion

par boucle grâce à un oligonucléotide de séquence  
 5' CTGGGGGCAACATCCAAGAATGCAAATAGAACTCTCC 3'.

b) Production de facteur VIIIΔIII dans des cellules CHO

5

Le vecteur pTG2307 ainsi obtenu est transféré dans les cellules CHODXB11 (Dr. L. Chasin Colombia University, New York, USA) par la technique de coprécipitation au phosphate de calcium et les cellules transformées sont 10 sélectionnées selon le procédé décrit dans la référence (14) mais en omettant l'aminoptérine. Les clones sont isolés après 1,5 à 2 mois de sélection puis testés pour leur production de facteur VIIIΔIII. Sur 14 clones testés, 7 produisent entre 1 et 6 U de Facteur VIII (activité coagulante) pour  $10^6$  cellules et pour 24 h. Ces niveaux de productions sont 12 fois supérieurs à ceux obtenus en utilisant le promoteur SV40 de l'Adénovirus 2 de pTG384 (construction pTG1509, décrit dans la demande française de brevet déjà citée, n° 8711415) (voir tableau II).

15

20

Tableau II : Production de facteur VIIIΔIII dans les cellules CHO en fonction du plasmide utilisé.

25

30

Plasmide	Activité coagulante U/ $10^6$ cellules x 24h
pTG1509	0.5
pTG2307	6.0

Exemple 3 : Expression du facteur von Willebrand à l'aide du plasmide pTG2330 (fig. 6)

Le plasmide de départ est le plasmide p delta plus précisément le grand fragment de p delta entre les sites EcoRI et SalI. Le promoteur et le leader du gène EIII sont introduits comme fragment EcoRI - HindIII de PTGEIII385 puis un adaptateur HindIII - BamHI provenant de M13TG131 (3) et le bloc du gène Néo fragment BamHI de M13TG1823 (précédemment décrit) lié au fragment BglII - SalI de pTG385. Le plasmide ainsi formé est désigné par pTG2323.

Pour obtenir le bloc DHFR, le plasmide pSBDH394 est digéré par HindIII et PvuII puis le bloc d'expression du facteur IX est substitué par le promoteur précoce de SV40, fragment HindII - HindIII de p delta. Le site HindIII est ensuite éliminé : le plasmide est digéré par HindIII, le site est rempli par l'ADN polymérase (fragment de Klenow) et le plasmide est religué.

Le bloc d'expression de DHFR est excisé par digestion SalI - SphI et cloné dans M13TG131 (3) coupé par ces mêmes enzymes donnant M13TG346. Par bloc DHFR, on entend donc une unité de transcription formée d'un bloc d'expression du gène DHFR sous le contrôle du promoteur précoce de SV40 privé de sa séquence activatrice. Le plasmide pTG2323 est ensuite digéré par EcoRI puis traité à la nuclease S1. Les extrémités sont remplies avec l'ADN polymérase (fragment de Klenow) puis ligées avec le bloc DHFR, fragment SmaI - SalI de M13TG346 (extrémité SalI remplie à la Klenow). La fusion du site SalI rempli avec le Klenow et le site EcoRI traité à la nuclease S1, régénère un site SalI.

Le bloc DHFR est intégré de sorte que les gènes EIII et DHFR sont transcrits dans des directions opposées (pTG2325). Un polylinker de séquence :

5' AGCTTCGCGAATTCTCGAG 3'  
3' AGCGCTTAAGAGCTCCTAG 5'

possédant les sites de restriction HindIII, NruI, EcoRI, XbaI et BamHI est introduit entre les sites HindIII et 10 BamHI de pTG2325, donnant pTG2328. Le polylinker permet d'insérer les ADNc à exprimer. Le vecteur d'expression du facteur von Willebrand (vWF) est obtenu en insérant dans le site EcoRI de pTG2328 l'ADNc sous forme de fragment EcoRI, donnant pTG2330.

15 L'ADNc codant pour vWF complet est préparé à partir d'ARNm extrait de poumon humain, un organe riche en tissu vasculaire et donc en cellules endothéliales qui sont reconnues comme site de synthèse de vWF. L'ADNc est 20 cloné dans un Bactériophage  $\lambda$ . Les clones obtenus sont criblés en utilisant des souches nucléotidiques (13).

Plusieurs clones ont été obtenus puis assemblés de façon à obtenir un clone d'ADNc codant pour la 25 protéine entière. Le plasmide contenant les séquences assemblées est constitué du vecteur pTZ18R (Pharmacia) contenant dans son site EcoRI l'ADNc du vWF du nucléotide -131 (considérant que l'ATG initiateur est à +1) au nucléotide 8557 (site SstI). Par construction l'ADNc comprend une extrémité polyG à son extrémité 5' et un 30 fragment SstI-EcoRI de M13TG131 (3) à son extrémité 3'.

Le vecteur pTG2330 contient deux unités de transcription : un bloc polycistronique vWF/néo et le 35 bloc DHFR. Il est possible par conséquent de sélectionner

les clones transformés en utilisant l'un des marqueurs néo ou DHFR, ou les deux simultanément.

Le vecteur pTG2330 a donc été transfété dans la lignée CHO-DHFR<sup>-</sup> DG44 (obtenue auprès du Dr. L. Chasin, Columbia University, New York, USA) en utilisant la méthode de précipitation au phosphate de calcium. Trois jours après le transfert les cellules sont trypsinées et réparties dans un milieu Alpha 1900 (GIBCO) supplémenté avec 10 % de sérum de veau foetal (SVF) et avec de l'antibiotique G418 à une concentration de 400 µg/ml ou dans un milieu Alpha 2000 supplémenté en SVF avec (à une concentration de 10 nM ou de 30 nM) ou sans méthotrexate. Après deux jours de sélection les cellules en G418 sont trypsinées et réparties soit dans un milieu Alpha 1900 supplémenté en SVF et contenant 600 µg/ml de G418, soit dans un milieu Alpha 2000 (GIBCO) supplémenté avec du SVF dialysé et contenant 200 ou 400 µg/ml de G418. Les cellules sont maintenues dans ces milieux pendant deux à trois semaines puis les clones ayant émergés sont testés pour leur production de vWF au moyen d'un test ELISA commercialisé par Diagnostic-Stago.

Des clones producteurs peuvent être obtenus dans toutes les conditions de sélection (Tableau I), les plus producteurs étant isolés en Alpha 2000 supplémenté avec du méthotrexate.

A ce stade, il peut être intéressant d'améliorer les conditions de transfert dans les cellules de façon à obtenir des clones capables de résister à des concentrations plus élevées de méthotrexate ou de pouvoir dériver à partir des clones déjà isolés des lignées résistant à des concentrations plus élevées de méthotrexate.

Tableau III Production des clones CHO-DG44 transformés par pTG2330 suivant les conditions de sélection

	Milieu de sélection	Nombre de clones obtenus ayant une production de vWF (en ng/10 <sup>6</sup> cellules/24h) de :				
		0-10	10-100	100-500	500-1000	1000-3000
5	Alpha 2000SVF	26	8	20	1	0
10	Alpha 2000 SVF Méthotrexate 10nM	5	0	3	3	4
15	Alpha 2000 SVF Méthotrexate 30nM	1	0	0	0	2
20						

REFERENCES

(1) RIGBY, P.W.J. Dans Genetic Engineering Vol. 3, édité par R. WILLIAMSON. Academic Press p88-141.

(2) ZIFF, F. (1980). Nature 287, 491-499.

(3) KIENY, M.P., LATHE, R. et LECOCQ, J.P. (1983). Gene 26, 91-99.

(4) TAHASASHI, K. et al. (1986). Nature 319, 121-126.

(5) MIYAMOTO, N.G. et al. (1984). Nucl. Acids Res. 12, 8779-8799.

(6) MESSING, J. et VIEIRA, J. (1982). Gene 19, 269-276.

(7) JAYE, M. et al. (1983). Nucl. Acids Res. 8, 2325-2335.

(8) LEE, F. et al. (1981). Nature 294, 228-232.

(9) SOUTHERN, P.J. et BERG, P. (1982). J. Mol. Appl. Genet. 1, 327-341.

(10) HALL, C.V. et al. (1983). J. Mol. Appl. Genet. 2, 101-109.

(11) GRAHAM, F.L. et A.J. VAN DER EB (1973). Virol. 52, 456-467.

(12) MILLER, J.H. (1972) in Experiments in molecular genetics. Cold Spring Harbor.

2642767

20

(13) PAVIRANI, A. et al. (1987). Biotechnology 5,  
389-392.

5 (14) MULLIGAN, R.C. et BERG, P. (1981). Proc. Natl. Acad.  
Sci. USA 78, 2072-2076.

(15) WU Q.Y. et al. (1988) Blood, 71, 1341-1346.

REVENDICATIONS

1. Vecteur d'expression d'une protéine hétérologue dans les cellules eucaryotes comprenant une séquence d'ADN codant pour la protéine hétérologue ainsi que les éléments assurant son expression, caractérisé en ce que les éléments assurant l'expression de la séquence codant pour la protéine hétérologue comprennent les éléments essentiels de la séquence leader et/ou de la séquence du promoteur du gène EIII de l'Adénovirus 2.
2. Vecteur selon la revendication 1 caractérisé en ce qu'il comprend tout ou partie de la séquence leader du gène EIII.
3. Vecteur selon la revendication 2 caractérisé en ce que la partie de la séquence leader du gène EIII est celle qui comporte l'intron (nucléotide 27981-28360).
4. Vecteur selon la revendication 3 caractérisé en ce que la partie de la séquence leader du gène EIII est celle qui commence à l'intron (nucléotide 27981) et s'étend jusqu'à l'ATG initiateur de la traduction.
5. Vecteur selon la revendication 1 caractérisé en ce qu'il comprend tout ou partie du promoteur du gène EIII.
6. Vecteur selon l'une des revendications 1 à 5, caractérisé en ce que il comporte tout ou partie du promoteur et la partie de la séquence leader qui comporte l'intron (nucléotide 27981-28360) de l'Adénovirus 2.

7. Vecteur selon la revendication 6 caractérisé en ce que il comporte en outre une séquence activatrice de transcription.

5       8. Vecteur selon l'une des revendications précédentes, caractérisé en ce que il comporte des éléments assurant son intégration dans le génome de la cellule hôte, de préférence la séquence d'ADN mitochondrial murin du plasmide de p delta.

10      9 . Vecteur selon l'une des revendications 1 à 8, caractérisé en ce que il comporte une séquence d'ADN supplémentaire en aval de la séquence d'ADN codant pour la protéine hétérologue.

15      10. Vecteur selon l'une des revendications précédentes caractérisé en ce qu'il comporte un gène de sélection.

20      11. Vecteur selon la revendication 10 caractérisé en ce que le gène de sélection se trouve en aval de la séquence d'ADN codant pour la protéine hétérologue.

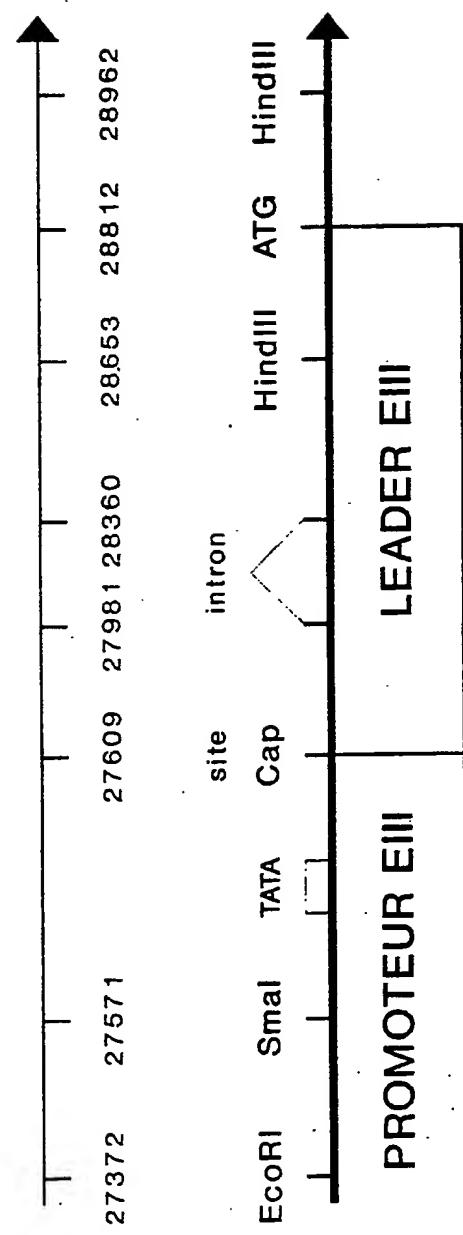
12. Vecteur selon la revendication 11 caractérisé en ce que le gène de sélection se trouve sur un bloc d'expression polycistronique, en aval de la séquence d'ADN codant pour la protéine hétérologue.

25      13. Vecteur selon la revendication 11 ou 12 caractérisé en ce qu'il s'agit du gène Néo.

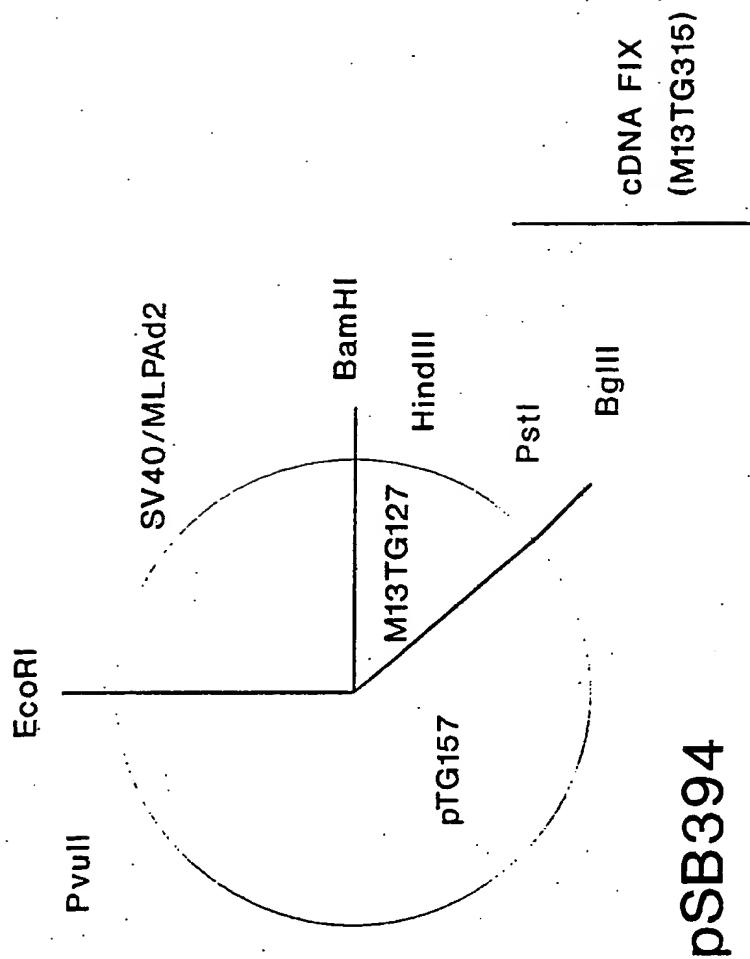
30      14. Vecteur selon l'une des revendications 10 à 13 caractérisé en ce que il comporte une unité de transcription d'un premier gène de sélection, et un bloc d'expression polycistronique pour la protéine hétérologue comportant un deuxième gène de sélection.

15. Vecteur selon la revendication 14 en ce que le premier gène de sélection est le gène DHFR.
16. Procédé de préparation de cellules eucaryotes afin d'obtenir des lignées cellulaires produisant une protéine hétérologue caractérisé en ce que on effectue la transformation de ces cellules par un vecteur selon l'une des revendications 1 à 15 et que l'on sélectionne les cellules exprimant ladite protéine hétérologue.  
5
17. Procédé selon la revendication 16 caractérisé en ce que la transfection est effectuée par la technique de coprécipitation au phosphate de calcium.  
10
18. Lignée cellulaire obtenue par la mise en oeuvre du procédé selon l'une des revendications 16 ou 17.  
15
19. Procédé de préparation d'une protéine hétérologue par des lignées cellulaires caractérisé en ce qu'on cultive une lignée cellulaire selon la revendication 18 sur un milieu de culture approprié et en ce qu'on récupère la protéine hétérologue à partir de la culture.  
20
20. Procédé selon la revendication 19 caractérisé en ce que la protéine hétérologue est choisie parmi les facteurs de la coagulation sanguine.  
25

**Fig. 1**



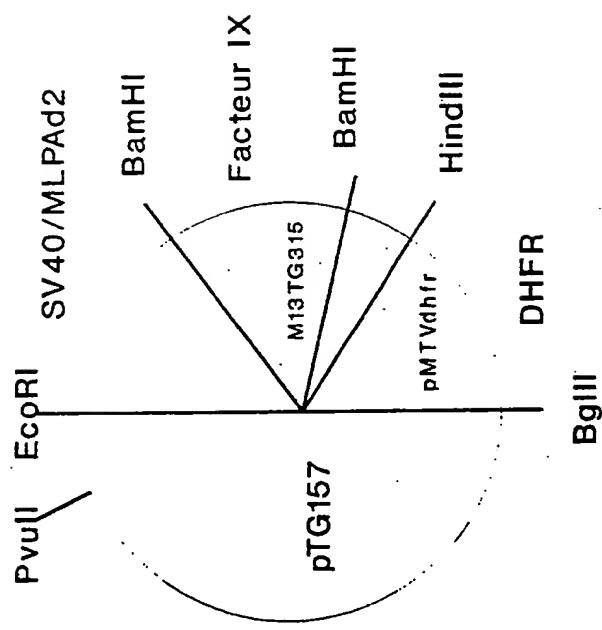
**Fig. 2**



II/6

2642767

**Fig. 3**

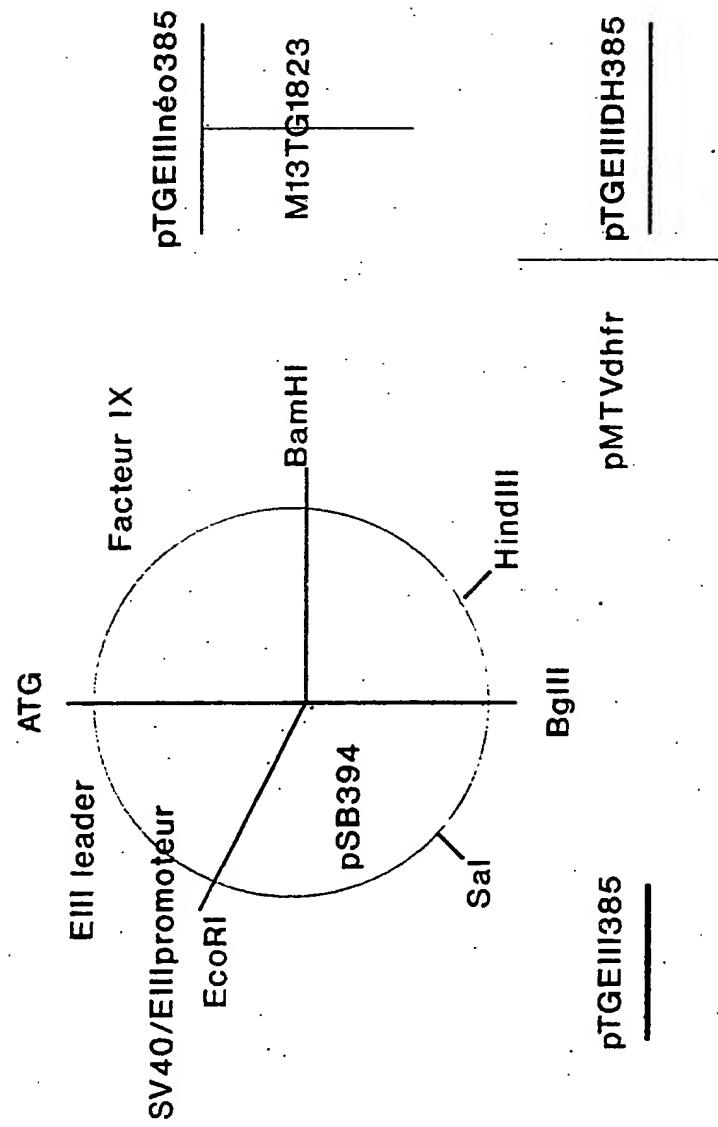


*III/6*

**2642767**

**pSBDH394**

Fig. 4



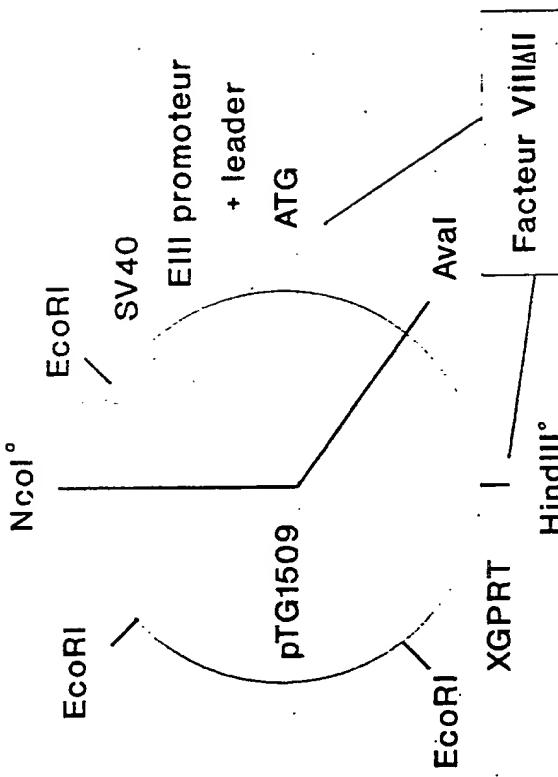
Iv/6

2642767

v/6

2642767

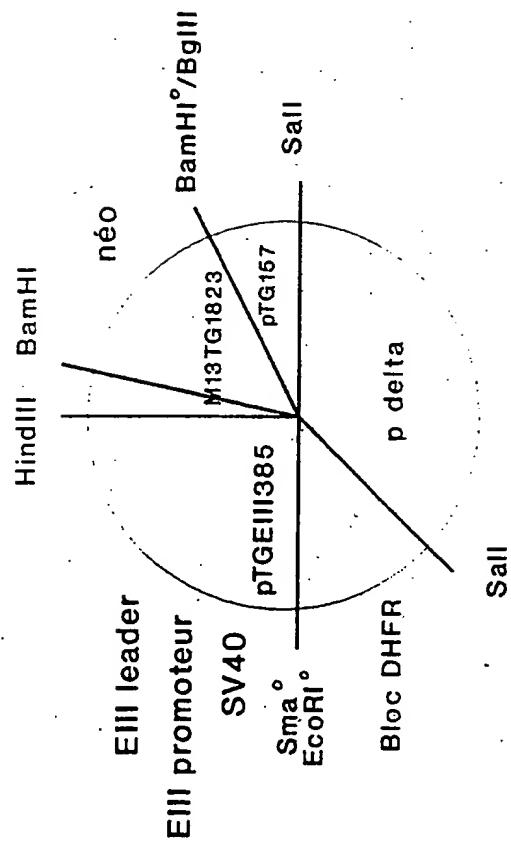
Fig. 5



pTG2307

**Fig. 6**

VWF



VI/6

2642767

**pTG2330**

(19) RÉPUBLIQUE FRANÇAISE  
INSTITUT NATIONAL  
DE LA PROPRIÉTÉ INDUSTRIELLE  
PARIS

(11) N° de publication :  
(à utiliser que pour les  
commandes de reproduction)

2 642 767

(21) N° d'enregistrement national :

89 00599

(51) Int Cl<sup>B</sup> : C 12 N 15/67, 15/86.

(12)

## DEMANDE DE BREVET D'INVENTION

A1

(22) Date de dépôt : 19 janvier 1989.

(71) Demandeur(s) : TRANSGENE S.A., Société anonyme. —  
FR.

(30) Priorité :

(72) Inventeur(s) : Henri de La Salle ; Thérèse Faure.

(43) Date de la mise à disposition du public de la  
demande : BOPI « Brevets » n° 32 du 10 août 1990.

(73) Titulaire(s) :

(60) Références à d'autres documents nationaux appa-  
rentés :

(74) Mandataire(s) : Cabinet Regimbeau, Martin, Schrimpf,  
Warcoin et Ahner.

(54) Vecteurs d'expression de protéines hétérologues dans les cellules eucaryotes, lignées cellulaires obtenues et  
procédé pour leur préparation.

(57) La présente invention concerne un vecteur d'expression  
d'une protéine hétérologue dans les cellules eucaryotes com-  
prenant une séquence d'ADN codant pour la protéine hétéro-  
logue ainsi que les éléments assurant son expression, caracté-  
risé en ce que les éléments assurant l'expression de la sé-  
quence codant pour la protéine hétérologue comprenant les  
éléments essentiels de la séquence leader et/ou de la sé-  
quence du promoteur du gène E1 de l'Adénovirus 2.

FR 2 642 767 - A1

La présente invention concerne de nouveaux vecteurs d'expression de protéines hétérologues dans des cellules eucaryotes, en particulier des cellules de mammifères.

5

Par protéine hétérologue, au sens de l'invention, on entend une protéine qui n'est pas exprimée naturellement par les cellules en cause, ou encore une protéine qui n'est pas nécessaire à la survie 10 de ces cellules.

De nombreux vecteurs faisant exprimer des gènes codant pour de telles protéines hétérologues dans des cellules de mammifères ont déjà été décrits. Le besoin de 15 disposer de vecteurs de plus en plus performants est permanent, notamment pour augmenter de façon sensible la quantité de produit synthétisé par le système hôte et permettre la sélection des cellules productrices et la préparation de lignées cellulaires stables produisant 20 ladite protéine en des quantités compatibles avec un rendement industriel. Différentes approches ont déjà été suggérées pour améliorer l'efficacité de ce type de vecteurs au niveau de la transcription comme au niveau de la traduction.

25

Ainsi par exemple de nombreux vecteurs utilisent en tant que promoteur de la transcription de la séquence d'ADN codant pour la protéine hétérologue, des séquences hybrides comprenant les séquences activatrices du virus 30 SV40 (répétitions de 72 paires de bases) associées au promoteur majeur tardif de l'Adénovirus 2. Une amélioration de l'efficacité de ce type de vecteur a déjà 35 été obtenue par différents moyens, en insérant des séquences d'ADN supplémentaires originaires de SV40 (répétitions de 21 paires de bases) et capables de lier

des facteurs de transcription ou encore des séquences du leader tripartite de l'Adénovirus 2 favorisant la traduction de l'ARN messager.

5 La présente invention vise à fournir de nouveaux vecteurs, utiles notamment pour l'expression de protéines complexes tels que les facteurs de coagulation, dans les cellules de mammifères.

10 Les vecteurs selon l'invention sont caractérisés en ce que les éléments assurant l'expression de la séquence codant pour la protéine hétérologue comprennent les éléments essentiels de la séquence leader et/ou de la séquence du promoteur du gène EIII de l'Adénovirus 2.

15 Par élément essentiel on entend la séquence assurant l'augmentation de la production de protéine et/ou le contrôle de l'expression de la séquence pour le promoteur.

20 Les adénovirus possèdent un ADN linéaire double brin qui dans le cas du sérotype 2 (Ad2) compte 35 kilobases. Les adénovirus ont cinq unités de transcription précoces ou plus dont celle du gène EIII et plusieurs unités de transcription tardives (1,2). La 25 carte de restriction des éléments amont du gène EIII est représentée sur la figure 1. Les coordonnées indiquées pour les nucléotides correspondent à celles obtenues à partir de la séquence Ad2 de la banque de donnée EMBL.

30 Parmi les éléments importants du gène EIII, on distingue le promoteur et la séquence leader, le promoteur allant jusqu'au site Cap, et la séquence leader allant depuis le site Cap jusqu'à l'ATG initiateur de la traduction.

Différentes variantes de l'invention peuvent être réalisées.

5 Dans une première variante, seule tout ou partie de la séquence leader du gène EIII, est présente sur le vecteur. De préférence, il s'agit d'une partie de la séquence qui comporte un intron, de sorte que le vecteur d'expression porte un intron dans la partie 5' non traduite de l'ARN messager (voir figure 1). Dans ce cas,

10 on peut utiliser comme promoteur, tous les promoteurs fonctionnels pour exprimer des gènes hétérologues dans les cellules de mammifères. On peut citer par exemple celui du cytomégalovirus ou celui du gène de la métallothionine.

15

Dans une seconde variante, le vecteur selon l'invention comprend tout ou partie du promoteur du gène EIII, c'est-à-dire la partie 5' des éléments de transcription allant jusqu'au site Cap.

20

Une variante préférée est celle où le vecteur comprend à la fois la partie de la séquence leader qui porte un intron et tout ou partie du promoteur du gène EIII de l'Adénovirus 2.

25

Il est bien sûr préférable d'avoir à la fois la séquence leader et le promoteur correspondant au même gène de l'Adénovirus 2. Il est avantageux, suivant une caractéristique supplémentaire de l'invention, d'ajouter 30 un activateur de transcription, en particulier les séquences activatrices de SV40 (par exemple les répétitions de 72 paires de bases, ou encore celle de 21 paires de bases ou une association des deux).

Les vecteurs selon l'invention sont en général des plasmides mais on peut également envisager des vecteurs víraux dans la mesure où il s'agit de virus dont le cycle comprend une étape dans laquelle le génome est sous forme d'ADN double brin comme par exemple les papillomavirus. Ils sont de préférence construits pour assurer l'intégration de la séquence codant pour la protéine hétérologue dans le génome de la souche hôte, de façon à obtenir des lignées cellulaires stables produisant ladite protéine. Afin d'obtenir des cellules comportant un grand nombre de copies du gène codant pour la protéine hétérologue, les vecteurs d'intégration peuvent comporter une séquence assurant l'intégration du vecteur en plusieurs centaines de copies comme par exemple le fragment d'ADN mitochondrial murin du plasmide p delta.

Dans certains cas, il peut être intéressant, pour augmenter la production de la protéine hétérologue, que le vecteur comporte une séquence d'ADN supplémentaire, en aval de la séquence d'ADN codant pour ladite protéine hétérologue.

Pour sélectionner les cellules transformées, le vecteur comporte de préférence un gène de sélection, par exemple le gène codant pour la xanthine-guanine phosphoribosyl-transférase (résistance à l'acide mycophénolique, XGPRT), le gène de résistance à la G418 (Néo), ou encore le gène dihydrofolate réductase (DHFR) qui confère la résistance au méthotrexate. On utilisera de préférence le gène Néo.

Le gène de sélection peut être porté par une unité de transcription indépendante, ou encore être placé directement en aval de la séquence d'ADN codant

pour la protéine hétérologue. Ce deuxième cas est particulièrement avantageux, puisqu'il permet d'obtenir un double effet : augmentation de la production et amélioration de l'efficacité de sélection.

5

Il peut aussi être avantageux de placer le gène de sélection dans un bloc d'expression polycistronique contenant un promoteur, une séquence d'ADN codant pour la protéine hétérologue, un gène de sélection, un intron, un site de polyadénylation. En effet, il a été démontré à plusieurs reprises que les ribosomes peuvent réinitier la traduction en aval du codon stop, et cette propriété est utilisée pour construire des plasmides contenant un bloc de sélection polycistronique. Comme gène de sélection, on peut employer ceux qui ont été cités plus haut, le gène de souris codant pour la dihydrofolate réductase (DHFR) et le gène bactérien de XGPRT.

Afin de pouvoir sélectionner plus facilement, après transformation des cellules de mammifères, celles qui expriment la protéine hétérologue en quantité, il est particulièrement avantageux d'ajouter à une unité polycistronique une unité de transcription d'un gène de sélection. De préférence, on utilisera un bloc d'expression du gène DHFR, par exemple sous le contrôle du promoteur précoce de SV40 privé de sa séquence activatrice.

L'invention a également pour objet un procédé de préparation de cellules eucaryotes, afin d'obtenir des lignées cellulaires produisant une protéine hétérologue selon lequel on effectue la transformation de ces cellules par un vecteur selon l'invention et on sélectionne les cellules exprimant lesdites protéines.

Parmi les méthodes de transformation utilisées, on peut citer en particulier la technique de coprécipitation au phosphate de calcium, qui sera mise en oeuvre dans les exemples ci-après.

5

Enfin, l'invention a pour objet les lignées cellulaires obtenues par la mise en oeuvre du procédé selon l'invention, en particulier des lignées CHO produisant du Facteur VIII ou du Facteur IX. En effet, 10 l'invention est applicable à la préparation des facteurs de la coagulation sanguine et concerne donc également un procédé pour leur préparation, selon lequel on cultive les cellules sur un milieu de culture approprié et on récupère la protéine obtenue à partir de la culture.

15

D'autres caractéristiques et avantages de l'invention apparaîtront au cours de la description détaillée suivante. Celle-ci indique quelques exemples de constructions selon l'invention et leur application à 20 l'expression de protéines complexes dans des cellules de mammifères.

Les figures suivantes illustrent l'invention :

25

- La figure 1 schématise la carte de restriction des éléments de transcription amont du gène EIII.

30

- La figure 2 représente schématiquement la structure du plasmide pSB394.

35

- La figure 3 représente schématiquement la structure du plasmide pSBDH394.

35

- La figure 4 représente schématiquement la structure des plasmides pTGEIII385, pTGEIIIDH385, pTGEIIINéo385.